Balance in competition in Dutch soccer

Koning, Ruud H.

Published in:
Journal of the Royal Statistical Society. Series D: The Statistician

DOI:
10.1111/1467-9884.00244

The Statistician (2000)
49, *Part* 3, *pp.* 419–431

Balance in competition in Dutch soccer

Ruud H. Koning

University of Groningen, the Netherlands

[Received August 1999. Revised April 2000]

Summary. We estimate an ordered probit model for soccer results in the Netherlands. The result of a game is assumed to be determined by home ground advantage and differences in quality between the opposing teams. The parameters of the model are used to assess whether the balance in competition in Dutch professional soccer has changed over time. Contrary to popular belief, we find that the balance has not changed much since the mid-1970s.

Keywords: Balance in competition; Home advantage; Ordered probit model; Professional soccer

1. Introduction

Professional soccer is a big business today. Broadcasting rights for 4 years have been sold in England for approximately US $1 billion, a typical sponsor contract for a top European team is valued at US $6 million (yearly) and the annual salary of a top striker is rumoured to be US $6 million. Demand, as measured by attendances in stadiums or numbers of spectators watching live broadcasts, has increased as well in recent years. In other words, the soccer business is becoming a major amusement industry (Economist, 1997).

The two main reasons for interest in a particular soccer game, or in any sports contest, are the quality of the play in absolute terms and the uncertainty about the outcome. The home games of weak teams are usually sold out when the top teams visit. Strong teams tend to win their games but sometimes they are taken by surprise by a weaker team. Most soccer *aficionados* can recall stories about a leader of a league who lost unexpectedly against a team that was at the bottom of the league. In fact, soccer results are more random than the results of games in other sports as only a few goals are scored each game and chance may be quite influential in determining the outcome. (The relationship between predictability and scores in different sports has been examined by Stefani (1983).)

Schemes such as sponsorship contracts, proceeds from the lucrative Champions League competition (a European competition in which only a selected number of teams participate), merchandizing and television rights allow wealthy teams to lure players away from poorer teams, even to sit on the substitutes' bench. As a result, weak teams are concerned that an increasing inequality in the income distribution of clubs leads to a decrease in the odds of beating strong teams. Poorer teams used to receive revenues from transferring players with great talent to top teams, but this source of income has vanished after the Bosman ruling. (According to the Bosman ruling by the European Court of Justice, a soccer player from the European Community is a free agent after his contract has expired.) This income could be used to improve training facilities for

Address for correspondence: Ruud H. Koning, Department of Econometrics, University of Groningen, PO Box 800, 9700 AV Groningen, the Netherlands.
E-mail: r.h.koning@eco.rug.nl

 0039–0526/00/49419

weaker teams or to increase the quality of the team by hiring players. The increased demand for top players across Europe has sent salaries sky high with obvious repercussions for the salary demands of mediocre players in average teams. These developments may cause a breakdown in balance in competition between teams and hence may decrease interest in soccer in the long run. Some weaker teams use these arguments to call for a redistribution of the proceeds of the sale of television rights (both of the national competitions and of the Champions League).

This paper examines the development of the balance in competition in Dutch professional soccer. Our aims are modest: we shall measure the balance in various ways, and we shall discuss its development over time. The structure of the paper is as follows. Section 2 discusses some theory on balance in competition. In Section 3 we develop a simple statistical model that can be used to analyse soccer results. The balance in competition and its evolution over time are discussed in Section 4. We end with some conclusions and directions for further research in Section 5.

2. Theory

Sports contests are interesting when there is not much difference in the quality of the contenders. As Quirk and Fort (1992), page 243, put it:

> 'One of the key ingredients of the demand by fans for team sports is the excitement generated because of uncertainty of outcome of league games. ... In order to maintain fan interest, a sports league has to ensure that teams do not get too strong or too weak relative to one another so that uncertainty of outcome is preserved.'

In fact, this is cited as the reason why some sports organizations in the USA are exempted from anti-trust regulation. Two teams engage in a joint production when they play a game. The outcome and the quality of the game are the good that is sold to the public. The public is worse off when the outcome of a game is easily predicted than if the game is tight. Therefore, collusion between teams to increase the quality of the game may be in the public's interest. This view neglects the absolute quality of the play. In fact, one of the important instruments to maintain balance in competitions in the USA is the inverse draft system, where lower ranked teams can pick talented new players before higher ranked teams can. In soccer leagues, there is no such balancing regulation.

According to the view cited above, an important task for sport bodies like the Union of European Football Associations or the Dutch Soccer Association is to maintain balance in competition because it is needed to ensure long-term interest in the league. The instruments that are available to achieve balance are limited, however. In the Netherlands, a court has decided in a preliminary ruling that individual teams are the owners of the broadcasting rights and not the organizing body. Each team can therefore sell its broadcasting rights individually and take the proceeds of this transaction. An implicit subsidy from wealthy teams to poor teams by the organizing body, to maintain balance, is no longer possible. Moreover, in contrast with baseball and football in the USA, gate receipts are not split between both teams. This may favour teams with big stadiums, even though a completely balanced competition is played (each pair of teams meets twice, once at each venue). In addition, there are no salary caps, either for the teams in total or for individual contracts in European soccer.

Balance in competition and regulations that intended to change it have been studied in the American context but not in the context of European soccer. Neumann and Tamura (1996) studied the balance in the National Football League in the USA. It is measured as the spread of parameters in a non-linear regression model. These parameters capture the quality of the teams. Bennett and

Fizel (1995) examined the effect of telecast deregulation on balance in competition in US college football. They measured it by comparing actual performances in a league with the performances that would be found if all teams were of equal strength (an approach developed by Noll (1991) and Scully (1989)). Empirical results have also been published in Quirk and Fort (1992). They measured the long-term development in five American professional sports leagues. They measured the balance by comparing the percentages of wins or losses for each league for each year with the percentages we would expect to find if all teams were equally strong. Each of the five leagues that they analysed showed a significant imbalance, though the imbalance in both baseball leagues has been decreasing in the last 20 or 30 years.

In the American literature, balance is usually defined as a win percentage of 50%. Such a definition is not useful for soccer, because of the prevalence of draws (unless we consider a draw to be a half-win). Draws are quite common in soccer. Over the period from 1956–1957 to 1996–1997, 26% of all league games ended in a draw, 48% ended in a win by the home team and the remaining 26% ended in a win for the away team. We define a soccer league to be in perfect balance for a certain year if the probability that any team wins a home game does not vary with the opponent or with the team. We assume that the home ground advantage is equal for each team, so in a balanced competition two teams would have equal probability of winning if the game were played on a neutral ground. In a balanced competition the probability that a team wins its home game may exceed the probability of a loss of a home game because of the home advantage. This definition allows for home advantage that changes over time, while the league is still in complete balance.

3. A model to analyse soccer results

3.1. *The statistical model*

In this section we propose a simple statistical model to analyse the outcome of soccer games. The model is an extension of the model of Neumann and Tamura (1996) in that we allow for an advantage for the home team. The strength of team i in the league is measured by a single parameter α_i. This parameter is independent of the opponent and venue of the game, and it is assumed to be constant during the season. If we assume that team i plays at home and team j is the away team, the difference in strength is $\alpha_i - \alpha_j$. To allow for unmeasured characteristics (i.e. those not captured by α), chance events during a game that influence the score etc., we assume that the outcome of the game is determined by the random variable D^*_{ij}:

$$D^*_{ij} = \alpha_i - \alpha_j + h_{ij} + \eta_{ij}, \qquad i, j = 1, \ldots, 18, \quad j \neq i. \tag{1}$$

In equation (1), h_{ij} is the home ground advantage of team i over team j which is assumed to be normally distributed with mean h. η_{ij} is a mean 0 random variable that captures other determinants of the result of the game. If D^*_{ij} is positive, team i is stronger than j, and D^*_{ij} is negative if team j is stronger than i. We do not observe the actual difference in strength; we only observe the outcome of the game. In fact, we observe whether team i has won, has played a draw or lost against team j. The latent difference in strength is transformed into an observed outcome of the game by

$$D_{ij} = \begin{cases} 1 & D^*_{ij} > c'_2, \\ 0 & c'_1 < D^*_{ij} \leqslant c'_2, \\ -1 & D^*_{ij} \leqslant c'_1 \end{cases} \tag{2}$$

with $D_{ij} = 1$ if team i wins, $D_{ij} = 0$ if team i plays a draw and $D_{ij} = -1$ if team j (the away team) wins the game. If we assume that h_{ij} and η_{ij} in equation (1) are independent normally distributed ($\epsilon_{ij} = h_{ij} + \eta_{ij} \sim \mathcal{N}(h, \sigma^2)$), then the probabilities of the possible outcomes of a game are

$$\begin{aligned} \Pr(D_{ij} = 1) &= 1 - \Phi\{(c_2 - \alpha_i + \alpha_j)/\sigma\}, \\ \Pr(D_{ij} = 0) &= \Phi\{(c_2 - \alpha_i + \alpha_j)/\sigma\} - \Phi\{(c_1 - \alpha_i + \alpha_j)/\sigma\}, \\ \Pr(D_{ij} = -1) &= \Phi\{(c_1 - \alpha_i + \alpha_j)/\sigma\} \end{aligned} \tag{3}$$

with $\Phi(\cdot)$ the standard normal distribution function and $c_1 = c_1' - h$ and $c_2 = c_2' - h$.

The statistical model in equation (2) allows for a constant home ground advantage. Consider two (hypothetical) teams of equal strength so that $\alpha_i - \alpha_j = 0$. The probability that the home team wins is $1 - \Phi(c_2/\sigma)$ and the probability that the home team loses is $\Phi(c_1/\sigma)$. These two probabilities are not constrained to be equal. In fact, we would expect that $\Phi(c_1/\sigma) < 1 - \Phi(c_2/\sigma)$ and this is confirmed by the results of our estimation. The existence of a home advantage can be examined formally by testing the hypothesis $c_1 = -c_2$.

It is not possible to identify all the parameters of this model. First, we need to fix the location of the quality parameters α. We impose the identifying restriction $\Sigma_i \alpha_i = 0$ so that the parameters α can be interpreted as deviations from a hypothetical average team with quality 0. A positive α_i implies that the quality of team i is better than average; a negative α_i implies the opposite. In addition, we fix the scale of the model by imposing the standard normalization $\sigma^2 = 1$.

Model (3) resembles models used by Clarke and Norman (1995), Stefani (1980) and Kuk (1995) to model soccer results. In Stefani (1980) and Clarke and Norman (1995), the dependent variable is the goal difference, and their models are estimated by least squares techniques. By using the least squares method, the dependent variable is assumed to follow a normal distribution. However, the observed goal difference takes only integer values with a limited range. They allow home ground advantage to vary between teams. The model in Kuk (1995) resembles the model above in that the dependent variable is the result of the game (and not the goal difference), and that an ordered probit model is used to derive the probability that a game is won, drawn or lost. In his model, the quality of a team differs between home and away games, and the home advantage varies between teams and over games. He estimated his model by using methods of moments using only the final ranking at the end of the season. An alternative approach could be based on Poisson-like models for the exact score in a game; see for instance Maher (1982), Dixon and Coles (1997) and Dixon and Robinson (1997). The reason that we prefer the ordered probit model is the simplicity of model (2): the quality of each individual team is captured by a single parameter. Moreover, this model allows for a simple separation between the measurement of quality of the teams and home advantage. In addition, the values that the dependent variable take are consistent with the stochastic specification of the model. Poisson-like models are usually more complex and have more parameters. For instance, in Maher (1982) at least two parameters per team had to be estimated and these parameters are difficult to interpret. Maher also assumed that the numbers of goals scored by the home team and the away team in a particular game are statistically independent. This may be too strong an assumption. He also assumed that games are mutually independent, an assumption that we shall make as well.

3.2. Description of the data and estimation results

The parameters were estimated by using the complete history of the Premier League of professional soccer in the Netherlands. (The data were obtained from Michael Koolhaas (private correspondence) and `http://www.noord.bart.nl/~kammenga/soccer`.) Organization of the competition as we know it today was introduced in the 1955–1956 season. In the 1962–1963, 1963–1964, 1964–1965 and 1965–1966 seasons only 16 teams participated in the Premier League. In all the other seasons 18 teams participated. Rules for relegation to the first division have changed over time. In the last few seasons the team finishing last was relegated automatically to the first division. The teams ranked 16 and 17 had to play additional games against teams in the first division. However, in earlier seasons, the teams finishing in the 17th and 18th places were relegated without having to play additional games. In total, 54 different clubs have played in the Premier League since the start of the competition in 1955–1956. Each year a couple of new teams entered the competition, either because of mergers or because of promotion. (A list of all the relevant mergers is given in Appendix A.) In each season, any combination of two teams meet twice: once at each venue. Therefore, a competition with 18 teams consists of 306 games; in total the data set comprises 12155 games. However, data from only 179 games are recorded from the 1996–1997 season.

We begin by estimating the parameters α_i, c_1 and c_2. It is assumed that the quality of the teams (measured by α_i) and the home advantage (measured by c_1 and c_2) were constant during the history of professional soccer. We use these results to calculate an all-time ranking of Dutch soccer teams by pooling the data over all seasons. The parameters were estimated by maximization of the log-likelihood function

$$l(\theta) = \sum_{\tau} \sum_{(i,j)\in\mathscr{T}_\tau} [I_{(D_{ij}=1)} \ln\{1 - \Phi(c_2 - \alpha_i + \alpha_j)\} + I_{(D_{ij}=0)} \ln\{\Phi(c_2 - \alpha_i + \alpha_j) - \Phi(c_1 - \alpha_i + \alpha_j)\} + I_{(D_{ij}=-1)} \ln\{\Phi(c_1 - \alpha_i + \alpha_j)\}]. \quad (4)$$

In this equation, τ is the index indicating the season and $\mathscr{T}_\tau$ is the index set of teams playing in the Premier League in season τ. The point estimates and their standard errors are given in Table 3 in Appendix A. The ordering of the parameters α_i indicates an all-time ranking. The best three teams have been Ajax (0.963), Feyenoord (0.750) and PSV Eindhoven (0.735), and the teams performing least well have been Dordrecht (−0.581), Fortuna SC (−0.498) and SVV (−0.433). This ranking is not necessarily equal to the standard ranking in which two or three points are awarded for each win and one for a draw. Teams that are relegated during some seasons do not earn any points in this ranking and would be at the bottom of the standard ranking. In our approach, there are no observations on a team if it does not participate in the Premier League during a season. Hence, the estimated α_i of a team that has participated in the Premier League for two seasons only can exceed the estimated α_i of a team that has played in the Premier League for, say, three seasons, if that first team played well during these two seasons. Indeed, we find that the teams with least points are not those with the smallest α_i: these are Fortuna SC, SHS and Alkmaar.

Second, we estimate the parameters for each year. This approach allows for variation in team-specific quality over time. Fig. 1 presents estimates of the home advantage. Detailed information on the individual estimates for each year is available on request. Home advantage is measured as the difference between the probability that the home team wins if both teams are of equal quality (i.e. $\alpha_i = \alpha_j$) and the probability that the away team wins. The circles depict the probability that the home team wins and the triangles the probability that the away team wins. If there had been no home advantage both sets would coincide (apart from sample variation). However, we see that there is a clear home advantage which has increased markedly during the second half of the 1960s to 33% in 1970–1971. Since the early 1970s, the probability that the home team wins against an opponent of equal strength is approximately 45–50%. The corresponding probability for the away team appears to have increased since then from approximately 15% to 20%. Therefore, home

Fig. 1. Probabilities of winning of home (O) and away (△) games when the teams are of equal strength

advantage decreased over that period. Indeed, over the 1990–1996 period, home advantage averaged 21% compared with 31% in the period 1970–1975. Home advantage in the league varies from year to year. This was also found by Clarke and Norman (1995). A test of whether $c_1 = -c_2$ is rejected at any reasonable level of significance for all years.

In Table 1, we present some summary statistics of the estimation results for each year. For each season we list the strongest and the weakest teams, and the standard deviation of the estimated α_is. Instead of giving the point estimates for the α_i-parameters, which are difficult to interpret, we give a transformation of these estimates. If team i has quality α_i in a given year, then the probability that this team wins a home game against the hypothetical team with quality 0 is $1 - \Phi(\alpha_i - c_2)$. This probability is given in the third and fifth columns of Table 1, where $\pi_{(n)}$ denotes the probability that the worst team in a given year beats the average team in a home game. $\pi_{(1)}$ is calculated similarly for the best team in that year. We have set $c_2 = 0.060$, the value obtained when estimating the model for the whole sample period. Hence, the variation in the probabilities reflects variations in quality, not changes in home advantage.

Note that the maximum likelihood estimate for an α_i would diverge to $-\infty$ if a team loses all its games during a season or to ∞ if a team wins all its games. No such teams are to be found even though Ajax came close in the 1971–1972 season by losing only one game, drawing in a mere three games and winning all the other games.

3.3. Specification tests

The model used in the previous sections is estimated by using maximum likelihood. The parameters are estimated inconsistently if the distributional assumption of normality is not correct. We tested whether we should reject the assumption of normality against the more general alternative that the distribution of the error terms belongs to a member of the Pearson family of distributions (for details of the test statistic we refer to Glewwe (1997) and Weiss (1997)). Members of the Pearson family include the normal, t- and Γ-distributions. Essentially, we test

Table 1. Best and worst team for each year

Year	*Worst team*	$\pi_{(n)}$	*Best team*	$\pi_{(1)}$	*Standard deviation*
1956	eindhoven	0.259	ajax	0.714	0.320
1957	den bosch	0.260	sc enschede	0.679	0.273
1958	shs	0.212	sparta	0.755	0.412
1959	sittardia	0.196	ajax	0.735	0.363
1960	noad	0.212	feyenoord	0.776	0.386
1961	rapid jc	0.263	feyenoord	0.730	0.308
1962	volewijckers	0.112	psv	0.700	0.399
1963	blauw-wit	0.265	dws	0.706	0.362
1964	nac	0.335	feyenoord	0.747	0.270
1965	heracles	0.297	ajax	0.867	0.466
1966	willem ii	0.143	ajax	0.863	0.537
1967	sittardia	0.212	ajax	0.898	0.529
1968	fortuna sc	0.175	feyenoord	0.894	0.628
1969	svv	0.144	ajax	0.918	0.600
1970	az	0.128	feyenoord	0.904	0.739
1971	volendam	0.168	ajax	0.959	0.725
1972	den bosch	0.179	ajax	0.931	0.678
1973	groningen	0.233	feyenoord	0.865	0.595
1974	wageningen	0.261	psv	0.852	0.520
1975	excelsior	0.216	psv	0.810	0.501
1976	de graafschap	0.273	ajax	0.794	0.429
1977	telstar	0.165	psv	0.795	0.443
1978	vvv	0.134	ajax	0.841	0.539
1979	haarlem	0.323	ajax	0.749	0.352
1980	wageningen	0.257	az	0.888	0.486
1981	de graafschap	0.122	ajax	0.852	0.552
1982	nec	0.259	ajax	0.887	0.514
1983	ds79	0.163	feyenoord	0.860	0.560
1984	pec zwolle	0.200	ajax	0.813	0.446
1985	heracles	0.123	psv	0.896	0.550
1986	excelsior	0.223	psv	0.894	0.484
1987	ds79	0.138	psv	0.877	0.476
1988	veendam	0.306	psv	0.808	0.370
1989	haarlem	0.167	ajax	0.736	0.403
1990	nec	0.292	psv	0.812	0.440
1991	vvv	0.120	psv	0.875	0.576
1992	fortuna sc	0.252	feyenoord	0.792	0.467
1993	cambuur	0.233	ajax	0.820	0.409
1994	dordrecht	0.238	ajax	0.914	0.556
1995	go ahead eagles	0.197	ajax	0.872	0.550
1996	az	0.237	psv	0.795	0.422

whether the third moment of ϵ is 0 and the fourth moment of ϵ is 3. The test statistic was calculated for each year that the model was estimated. The null hypothesis of normality was not rejected in any season at a 5% level of significance. This conclusion is not at odds with the Poisson assumption that is often made when analysing soccer scores. The score difference is estimated by summing over a large number of Poisson-distributed scores, and hence in the limit it can be approximated by a normal distribution. Clarke and Norman (1995) also found that the residuals were approximately normally distributed when they estimated a model similar to equation (1) by least squares.

We tested whether or not restrictions on the parameters could be imposed. First, for each year all estimated α_i were found to be jointly significantly different from 0. Second, the hypothesis that the home advantage is constant over time had to be rejected. We also rejected the hypothesis that the quality of a given team does not vary over time.

If we assume that the quality parameters α_i remain constant over time, it is possible to test for variation in the variance of the error term in model (2). We imposed this restriction, and re-estimated the model for the period from 1991–1992 to 1996–1997 with unrestricted variances. We could not reject the null hypothesis that the variance of ϵ_{ij} is constant over time.

4. Empirical evidence of balance in competition

In this paper we measure the balance in competition in three different ways.

(a) If the standard deviation σ_P of the number of points in the final ranking of a competition is small, there is not much spread in the points gained at the end of the season and the competition has been tight.

(b) Since α_i is the extent to which team i is better than a hypothetical team with quality 0, it is natural to measure the balance in competition by the total deviation from average quality, $\Sigma_i \alpha_i^2$. This is proportional to the standard deviation σ_α of the quality parameters of the statistical model of the previous section. Again, if this number is small, the quality of the teams does not vary much.

(c) The concentration ratio CR_K is defined as the number of points obtained by the top K teams divided by the number of points that they could have gained. If there are J teams in a competition the team winning the competition could have obtained $2W(J-1)$ points where W is the number of points awarded for a game won. We denote the number of points obtained by the kth-best team by $P_{(k)}$. The concentration ratio is formally defined as

$$\mathrm{CR}_K = \sum_{k=1}^{K} P_{(k)} \Big/ KW(2J - K - 1), \tag{5}$$

the number of points actually obtained by the K best teams divided by the maximum number of points that they could have obtained. If the concentration ratio is high, the top K teams did not lose many points to weaker teams.

These three measures are to some extent ‘static’ as they refer to balance *within* a particular season. An advantage of the second measure compared with the first measure is that it does not require that the season be complete. The concentration ratio is not a measure of balance in the whole competition; it applies to the quality of the top teams. This measure is interesting though as it is commonly believed that the gap between top teams and the rest has increased over time. It is not possible to address this issue with the first two measures.

The first two measures are not completely equivalent: a crucial drawback of σ_P as a measure of balance in competition is that it is not invariant under changes in home advantage. As our statistical model separates home advantage and team quality, σ_α does not suffer from this drawback. Moreover, in the 1995–1996 season the number of points obtained for a win was raised from 2 to 3. In contrast with the standard deviation of the number of points, the standard deviation of the α_i is invariant to changes in the number of points awarded for a win or a draw. Finally, we can estimate the α_i and their standard deviations even if a season is not finished completely. Note that σ_P and σ_α may rise if the average scoring frequency of all teams increases. In this case, skilled teams accumulate more wins whereas less skilled teams accumulate more losses, leading to more dispersion in the point and quality distributions. This is of little practical importance as goal scoring does not vary much over the years.

Estimates of σ_P and σ_α are graphed in Fig. 2. The standard deviation of the α_i varies between

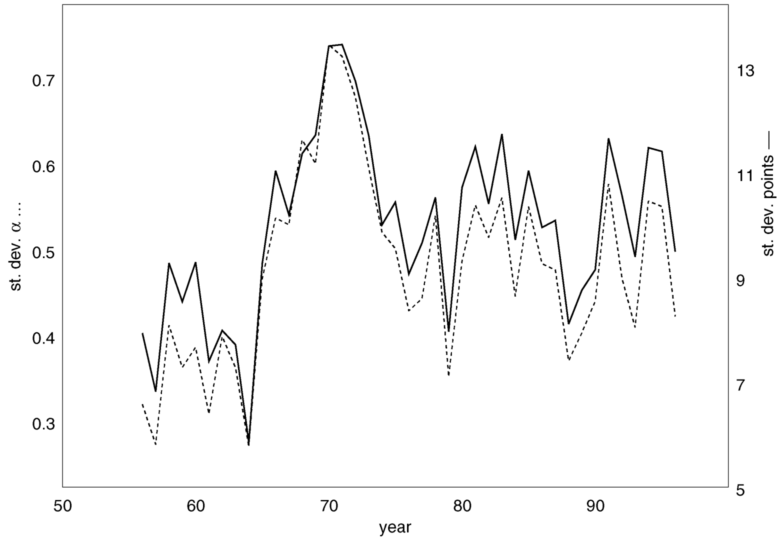

Fig. 2. Standard deviations of the estimated α_i (········) and of the numbers of points (———)

0.25 and 0.75, and the standard deviation of the number of points varies between 5.9 and 13.5. To interpret the level of σ_P and σ_α we have simulated two hypothetical competitions and calculated σ_P and σ_α for each simulated competition. In the first competition, each team is equally strong and has a 50% probability of winning a home game, a 25% probability of playing a draw and 25% of losing a home game. For this case, the average value of σ_P is 4.92, and the average value of σ_α is 0.20.

In the second simulation the probabilities are the same, except for three strong teams. These three strong teams win games against each of the remaining 15 weaker opponents with 90% probability (and a 5% probability of a draw and a 5% probability of a loss). In this second case, the average value of σ_P is 9.87, and the average value of σ_α is 0.46. Therefore, a competition with three strong teams and 15 identical weaker teams gives the same values for σ_P and σ_α as we find in Fig. 2.

The balance in competition has not changed systematically from the early start of professional soccer in 1955 until the mid-1960s. We then see a marked decrease between 1965 and 1970 followed by an increase in between 1970 and 1976. Coincidentally (or not) it was in the period 1966–1970 that Dutch professional soccer caught up with the best teams in Europe. Ajax was the first Dutch finalist in a European tournament in the spring of 1969 and Feyenoord was the first Dutch winner in a European tournament in 1970. Dutch (and European) soccer was dominated by Ajax from 1970 until 1973, a period when the balance in competition in Dutch soccer increased sharply.

From the mid-1970s to the end of the century there is no clear trend in the balance in competition. One year, competition is tighter than in another year, but no clear trends are discernible from the data. The spread from year to year is considerable; years with a tight competition are interspersed with years with a one-sided competition. This is especially noteworthy because it was feared in the early 1980s that competition would become less tight because of shirt sponsorship which in Dutch soccer has been allowed from the 1981–1982 season onwards. At first, it seemed that criticism of shirt sponsorship was justified, since in the 1981–

1982 season two teams could not find a sponsor at all. This was only temporary though: even amateur teams now have shirt sponsors. Some teams have had better sponsoring deals than others, resulting in a more unequal distribution of income. This increase in income inequality is claimed to lead to a decrease in balance in competition. However, we do not find any evidence for this hypothesis. Less successful teams use the same arguments as those used against shirt sponsorship to oppose television contracts that give a larger share of the revenue to more successful teams. According to them, an unequal distribution of television revenues will lead to an unequal distribution of quality and this leads to a decrease in general interest in soccer. However, the balance in competition has not decreased significantly since the introduction of shirt sponsorship. In fact, shirt sponsorship may have enabled most semiprofessional players to become full professionals and this may have increased the overall quality of soccer.

To examine the robustness of our results, we have estimated two other models. First, we estimated the quality parameters of the teams by using the model of Clarke and Norman (1995). In this model the dependent variable was the difference in goals scored, and the parameters were estimated by using least squares. In the second model, the ordered probit structure of equation (3) was retained, but two additional categories were added. The dependent variable now makes a distinction between a win (or loss) by a three-goal difference (or more). In both cases, the variation in the estimated quality parameters was examined, and it turned out that the results were very similar to those in Fig. 2. Hence, our conclusions about the development of the balance in competition are robust with respect to the statistical model used.

As a third indicator of the balance in competition, we look at the concentration ratios for the first and fourth place. The results are given in Fig. 3. Qualitatively, we see the same picture as in the previous graphs: until the mid-1960s the top team in each year captured only 75% of the number of points that it could have obtained at the end of the season. This percentage increases during the second half of the 1960s, to a maximum of 94% for the top team in 1971–1972. Then an increase in the balance of competition sets in, followed by an irregular period with no clear trends. The picture is slightly different for the top four teams: a slight upward trend of the

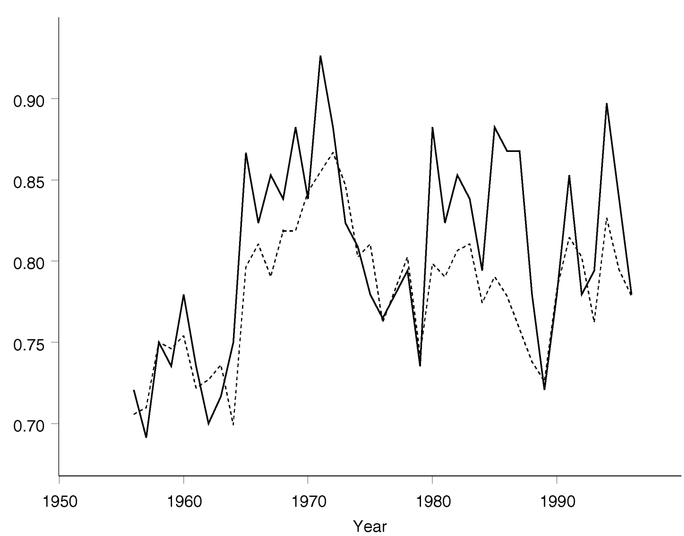

Fig. 3. Concentration ratios for the first and fourth places: ———, first place; ··········, fourth place

concentration ratio during the 1980s and 1990s is visible. (CR_4 exceeds CR_1 in a particular year if the number of points obtained by the teams that end second, third and fourth does not differ much from the number of points obtained by the team that ended first (i.e. if $P_{(2)} + P_{(3)} + P_{(4)} > (45/17)P_{(1)}$). This happens in 13 seasons.) Contrary to common opinion and despite the slight upward trend, the value of CR_4 is not high when compared with typical values encountered during the 1960s.

5. Conclusion

In this paper we have discussed the balance in competition in Dutch professional soccer. We used a simple model to estimate the quality of the teams participating in the Premier League. We find that the balance decreased markedly during the second half of the 1960s and increased during the first half of the 1970s but that there has been no clear trend since. We also find that the introduction of shirt sponsorship did not lead to a noticeable significant decrease in the balance. These conclusions are borne out by three different measures of balance and hence are not sensitive to model specification.

This paper provides only a starting-point for a more structural economic analysis of the balance in competition. The 'superstar' model of Rosen (1981) may provide insight into why an increase in income inequality that may have taken place did not lead to a decrease in the balance of competition. Another issue to be resolved is to examine whether the lack of recent trends in the balance of competition is specific to Dutch soccer or whether it is a more international phenomenon.

Acknowledgements

The author thanks Marco Haan, Peter Hopstaken, Bas van der Klaauw, Geert Ridder, seminar participants at the University of Mannheim, Concordia University, Queen's University and participants at the Statistische Dag and the European Economic Association meeting in Berlin for helpful discussions and comments. The exposition and the content of the paper have benefited from detailed comments by three referees.

Appendix A: Estimation results of the complete model

In this appendix we discuss the construction of our data set and give estimation results of the model with the α_i constant during the sample period. First, in Table 2 we give a list of mergers in Dutch professional soccer. (The list is based on information provided in Verkammen and Vermeer (1994).) Some teams have played under several names. For instance, FC Dordrecht changed its name in 1979 to DS '79 and in 1990 again to Dordrecht '90. In other cases, professional soccer teams merged with other professional soccer teams (for example, in 1991 Dordrecht '90 merged with SVV to form SVV/Dordrecht '90 which changed its name again in 1992, reverting to Dordrecht '90). We have treated each team that resulted from a merger as a new team, so we distinguish between DS '79 (a predecessor of Dordrecht '90 that played in the Premier League in 1987–1988) and Dordrecht '90 that resulted from a merger with SVV. In the same vein, FC Amsterdam before 1974 is considered to be a different team from FC Amsterdam after the year when it merged with De Volenwijckers.

In Table 3 we give the estimation results of the model estimated over the period 1956–1996 in which all parameters are constant over time. The number of cases is 12155, and the mean log-likelihood is -0.980507.

Table 2. Mergers in Dutch professional soccer†

Year	*New team*	*Merged teams*
1958	DWS	Amsterdam, *DWS*
1962	Roda JC	*Roda Sport*, Rapid JC
1963	Telstar	*Stormvogels*, *VSV*
1965	Twente	SC Enschede, *Enschedese Boys*
1967	AZ	Alkmaar, *Zaanstreek*
1967	Den Bosch	*Den Bosch*, *Wilhelmina*
1967	Xerxes/DHC	*DHC*, Xerxes
1968	Fortuna SC	Fortuna ’54, Sittardia
1970	Utrecht	Dos, Elinkwijk, *Velox*
1971	Den Haag	ADO, Holland Sport
1972	FC Amsterdam	DWS, Blauw Wit
1974	FC Amsterdam	FC Amsterdam, De Volenwijckers
1991	Dordrecht ’90	*Dordrecht ’90*, SVV

†Teams in italics have never played in the Premier League.

Table 3. Estimation results of the full model

Team	$\hat{\alpha}$	*Standard deviation*	*Team*	$\hat{\alpha}$	*Standard deviation*	*Team*	$\hat{\alpha}$	*Standard deviation*
ado	0.322	0.032	fortuna 54	0.159	0.029	shs	−0.546	0.174
ajax	0.951	0.031	fortuna sc	0.009	0.022	sittardia	−0.212	0.040
alkmaar	−0.219	0.043	go ahead eagles	0.057	0.028	sparta	0.262	0.029
amsterdam	0.034	0.033	graafschap	−0.046	0.030	svv	−0.443	0.047
az	0.215	0.032	groningen	0.123	0.021	telstar	−0.082	0.033
blauw-wit	0.023	0.031	haarlem	−0.032	0.032	twente	0.389	0.028
cambuur	−0.297	0.040	heerenveen	0.105	0.041	utrecht	0.127	0.023
de graafschap	−0.446	0.054	helmond	−0.446	0.044	veendam	−0.255	0.038
den bosch	−0.067	0.022	heracles	−0.192	0.051	vitesse	0.252	0.031
den haag	−0.002	0.025	holland	0.031	0.031	volendam	−0.050	0.021
dordrecht	−0.335	0.050	mvv	0.048	0.032	volewijckers	−0.353	0.035
dos	0.156	0.019	nac	0.045	0.033	vvv	0.004	0.021
ds79	−0.829	0.113	nec	−0.085	0.029	wageningen	−0.356	0.045
dws	0.141	0.023	noad	−0.235	0.040	willem ii	−0.047	0.022
eindhoven	−0.283	0.057	pec zwolle	−0.103	0.041	xerxes	0.212	0.085
elinkwijk	−0.136	0.051	psv	0.724	0.029	xerxes/dhc	0.272	0.116
excelsior	−0.236	0.039	rapid jc	0.111	0.020	c_1	−0.721	0.012
fc amsterdam 1	0.280	0.049	rkc	0.107	0.019	c_2	0.060	0.010
fc amsterdam 2	−0.152	0.026	roda jc	0.267	0.034			

References

Bennett, R. W. and Fizel, J. L. (1995) Telecast deregulation and competitive balance: NCAA Division I football. *Am. J. Econ. Sociol.*, **54**, 183–199.

Clarke, S. R. and Norman, J. M. (1995) Home ground advantage of individual clubs in English soccer. *Statistician*, **44**, 509–521.

Dixon, M. J. and Coles, S. G. (1997) Modelling association football scores and inefficiencies in the football betting market. *Appl. Statist.*, **46**, 265–280.

Dixon, M. J. and Robinson, M. E. (1997) A birth process for association football matches. Lancaster University, Lancaster.

Economist (1997) Golden goals. *Economist*, May 31st, 63–65.

Glewwe, P. (1997) A test of the normality assumption in the ordered probit model. *Econometr. Rev.*, **16**, 1–19.

Kuk, A. Y. C. (1995) Modelling paired comparison data with large numbers of draws and large variability of draw percentages among players. *Statistician*, **44**, 523–528.

Maher, M. J. (1982) Modelling association football scores. *Statist. Neerland.*, **36**, 109–118.

Neumann, G. R. and Tamura, R. F. (1996) Managing competition: the case of the national football league. University of Iowa, Iowa City.

Noll, R. G. (1991) Professional basketball: economic and business perspectives. In *The Business of Professional Sports* (eds P. D. Staudohar and J. A. Mongan), pp. 18–47. Chicago: University of Illinois Press.

Quirk, J. and Fort, R. D. (1992) *Pay Dirt*. Princeton: Princeton University Press.

Rosen, S. (1981) The economics of superstars. *Am. Econ. Rev.*, **71**, 845–858.

Scully, G. W. (1989) *The Business of Major League Baseball*. Chicago: University of Chicago Press.

Stefani, R. T. (1980) Improved least squares football, basketball, and soccer predictions. *IEEE Trans. Syst. Man Cyber.*, **10**, 116–123.

———(1983) Observed betting tendencies and suggested betting strategies for European football pools. *Statistician*, **32**, 319–329.

Verkammen, M. and Vermeer, E. (1994) *Om 't Spel en de Knikkers*. Bilthoven: Mundt.

Weiss, A. A. (1997) Specification tests in ordered logit and probit models. *Econometr. Rev.*, **16**, 361–391.